Université de Montréal
École des Hautes Études Commerciales
École Polytechnique

Université de Montréal
C.P. 6128, Succursale Centre-ville
Montréal (Québec) Canada H3C 3J7

Telephone: +1-514-343-7575
Telecopieur (Fax): +1-514-343-7121
crt@crt.umontreal.ca
www.crt.umontreal.ca/CRT/

CRT-2002-15

Centre de recherche sur les transports (C.R.T.)
Centre for Research on Transportation

AN AFFINE-INVARIANT GENERALIZATION OF THE WILCOXON SIGNED-RANK TEST FOR THE BIVARIATE LOCATION PROBLEM

by

Denis Larocque
Serge Tardif
Constance van Eeden

June 2002

PUBLICATION
CRT-2002-15

AN AFFINE-INVARIANT GENERALIZATION OF THE WILCOXON SIGNED-RANK TEST FOR THE BIVARIATE LOCATION PROBLEM*

by

Denis Larocque[1]
Serge Tardif[2]
Constance van Eeden[3]

Dedicated to the memory of Serge Tardif.

* This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR).

[1] Department of Management Sciences, École des Hautes Études Commerciales, 3000 chemin de la Côte-Sainte-Catherine, Montréal H3T 2A7 Canada, and Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal H3C 3J7 Canada

[2] Serge Tardif died on September 26, 1998 after an earlier version of this paper had been prepared. The remaining authors take full responsibility for the present version.

[3] Department of Statistics, The University of British Columbia, Vancouver, Canada V6T 1Z2

Dépôt légal – Bibliothèque nationale du Québec
Bibliothèque nationale du Canada, 2002

Abstract

A test extending the univariate Wilcoxon signed-rank test to the bivariate location problem is proposed. Two versions of the test statistic are given an each of them is affine-invariant. The first version is a conditionally distribution-free test that can be used with any sample size. The second version can be used for larger sample sizes and has a limiting χ_2^2 distribution under the null hypothesis. The relationship with a test proposed by Jan and Randles (1994) is investigated. It is shown that the Pitman efficiency of this test relative to the new test is equal to one for elliptical distributions but that the two tests are not necessarily equivalent for non-elliptical distributions. These facts are also demonstrated empirically with a simulation study. The new test has the advantage of not requiring the assumption of elliptical symmetry that is needed to perform the asymptotic version of the Jan and Randles test.

Key words : bivariate location, affine-invariance, Wilcoxon signed-rank test, conditionally distribution-fee.

Résumé

Dans cet article, nous proposons un test de position bivarié qui généralise le test de Wilcoxon des rangs signés univariés. Deux versions du test sont données et chacune d'elles est "affine-invariant". La première version du test mène à un test conditionnellement " distribution-free " qui peut être utilisé, peu importe la taille de l'échantillon. La deuxième version peut être utilisée pour des échantillons de plus grandes tailles et sa loi limite est la loi du khi-deux. Nous étudions le lien entre notre test et celui proposé par Jan et Randles (1994). Nous montrons que ces deux tests sont asymptotiquement équivalents en terme d'efficacité asymptotique de Pitman pour les distributions elliptiques mais que ce n'est pas nécessairement le cas pour les distributions non elliptiques. Ceci est également démontré empiriquement à l'aide d'une simulation. Contrairement à la version asymptotique du test de Jan et Randles, le nouveau test est valide sans qu'il ne soit nécessaire de supposer que les observations sont de loi elliptique.

Mots-clefs : problème de position bivarié, " affine-invariance ", test de Wilcoxon des rangs signés, test conditionnellement " distribution-free ".

1 Introduction

Let $\tilde{\mathbf{X}}_1 = (X_1, Y_1), \ldots, \tilde{\mathbf{X}}_n = (X_n, Y_n)$ be independently and identically distributed (iid) bivariate random vectors. It is assumed that $\tilde{\mathbf{X}}_1$ is distributed symmetrically around $\boldsymbol{\Delta} = (\delta_1, \delta_2)$, that is, $\tilde{\mathbf{X}}_1 - \boldsymbol{\Delta}$ and $\boldsymbol{\Delta} - \tilde{\mathbf{X}}_1$ are identically distributed. The problem is to test the hypotheses

$$H_0 : \boldsymbol{\Delta} = \mathbf{0} \quad \text{and} \quad H_1 : \boldsymbol{\Delta} \neq \mathbf{0}. \tag{1}$$

Assuming normality of the random vectors, the classic procedure for this problem is Hotelling's T^2 test. An important feature of this test statistic is that of affine-invariance, that is, its value remains unchanged after a nonsingular linear transformation is applied to the observations. This property ensures that the decision made will not be affected by the units of measurement and/or the choice of axis.

Many bivariate generalizations of the univariate signed-rank test have been proposed as competitors to Hotelling's T^2 test, Hettmansperger & McKean (1998 Chapter 6). The problem is that some of the obvious generalizations are not affine-invariant. The coordinate-wise signed-rank test of Bennett (1964) is scale-invariant but not rotation-invariant. It has been shown in numerous empirical studies, e.g. Larocque, Tardif & van Eeden (2002), that the lack of rotation-invariance of the tests based on this coordinate-wise approach can cause their power to decrease substantially for certain correlation structures. Hössjer & Croux (1995) and Möttönen & Oja (1995) introduced spatial signed-rank tests that are rotation-invariant but not scale-invariant. Likewise, the power of those tests can suffer when the variables are measured on different scales; Larocque *et al.* (2002). Examples of affine-invariant bivariate signed-rank tests are the one proposed by Brown & Hettmansperger (1987), based on Oja's measure of scatter, and the ones proposed by Peters & Randles (1990) and Jan & Randles (1994), based on the concept of interdirections. For the last two, the assumption of elliptical symmetry is required for the derivation of the asymptotic null distribution. For spherical distributions, the Peters and Randles and Hössjer and Croux tests are asymptotically equivalent and the same is true for the Jan and Randles and the Möttönen and Oja tests.

103-03064/

The purpose of this paper is to introduce a bivariate affine-invariant generalization of the Wilcoxon signed-rank test. The advantage of this test over the one of Jan and Randles is that its asymptotic null distribution holds without the assumption of elliptical symmetry. It also requires less computations to perform. Furthermore, unlike the spatial signed-rank test of Möttönen and Oja, which is only rotation-invariant, this test is affine-invariant. In Section 2, we present the statistic and some of its properties. In Section 3, we discuss of Pitman efficiency, about the relation with the work of Jan and Randles and present the results of a small simulation study. Concluding remarks and a numerical example are given in Section 4. The Appendix contains some of the proofs.

2 The test statistic and its properties

Let, for each pair (j,k), $1 \leq j < k \leq n$,

$$\theta_{jk} = -\arctan\left(\frac{X_j + X_k}{Y_j + Y_k}\right) + \frac{\pi}{2} \tag{2}$$

be the angle ($\in [0,\pi)$) between the vector $(X_j + X_k, Y_j + Y_k)$ and the X-axis. The test will be based on the two statistics

$$\begin{aligned} A_n &= \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} s(Y_j + Y_k) \cos\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right) \\ B_n &= \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} s(Y_j + Y_k) \sin\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right), \end{aligned} \tag{3}$$

where R_{jk} is the rank of θ_{jk} among the $n(n-1)/2$ θ_{rt}s and where $s(u) = 1$, 0 or -1 as $u > 0$, $= 0$ or < 0 is the sign function.

One motivation for these statistics relies on the fact that they are closely related to the L_2-norm of the Wilcoxon signed-rank statistic computed using the projection of the points (X_i, Y_i) on lines passing through the origin. Indeed, let $P_j(\theta) = X_j \cos(\theta) + Y_j \sin(\theta), j = 1,\ldots,n$, be the projection of (X_j, Y_j) on the directed line with angle θ. Then,

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} s(P_i(\theta) + P_j(\theta))$$

is the Wilcoxon signed-rank statistic for the projected observations. Define, for $1 \leq i < j \leq n$, $Z_{ij} = \theta_{ij} - \pi/2$ and let F_Z be the cumulative distribution function (cdf) of Z_{12}. If F_Z was known, an intuitively appealing statistic for testing the hypothesis H_0 would be

$$W_p = \int_{-\pi/2}^{\pi/2} \left(\frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} s(P_i(\theta) + P_j(\theta)) \right)^2 dF_z(\theta).$$

By straightforward algebra and by replacing F_Z by the empirical cdf of $Z_{12}, Z_{13}, \ldots, Z_{(n-1)n}$, W_p becomes

$$\frac{1}{\binom{n}{2}^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=1}^{n-1} \sum_{l=k+1}^{n} s(Y_i + Y_j) s(Y_k + Y_l) \left(1 - \frac{2}{\binom{n}{2}} |R_{ij} - R_{kl}| \right).$$

Now, this statistic can be viewed as a member of the following family:

$$\frac{1}{\binom{n}{2}^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=1}^{n-1} \sum_{l=k+1}^{n} s(Y_i + Y_j) s(Y_k + Y_l) \phi \left(\frac{|R_{ij} - R_{kl}|}{\binom{n}{2}} \right),$$

where $\phi : [0,1] \to [-\infty, \infty]$ is a function satisfying $\phi(u) = -\phi(1-u)$ for all $u \in [0,1]$. If we select $\phi(u) = \cos(\pi u)$, we get the statistic

$$C_n = A_n^2 + B_n^2. \tag{4}$$

As shown in the Appendix, see the proof of Theorem 2.1 a), this statistic is affine-invariant. Although it is not distribution-free, it can be used to perform a conditionally distribution-free test. The way to perform such a test is explained in Jan & Randles (1994). Let $\boldsymbol{\Gamma} = (\gamma_1, \ldots, \gamma_n)$ where γ_i =−1 or 1. Let $\tilde{\mathbf{X}}_{\boldsymbol{\Gamma}} = (\gamma_1 \tilde{\mathbf{X}}_1, \ldots, \gamma_n \tilde{\mathbf{X}}_n)$. Under H_0, the 2^n sets of observations $\tilde{\mathbf{X}}_{\boldsymbol{\Gamma}}$ are equally likely. The conditional distribution of C_n can be obtained by considering its distribution over those 2^n sets of observations. The conditional P-value is then simply computed using this distribution as reference. More specifically, let $C_n(\tilde{\mathbf{X}}_{\boldsymbol{\Gamma}})$ be the value of C_n using the observations $\tilde{\mathbf{X}}_{\boldsymbol{\Gamma}}$. The conditional P-value is the proportion of $C_n(\tilde{\mathbf{X}}_{\boldsymbol{\Gamma}})$, among the 2^n, that is greater than or equal to the value of C_n obtained from the original observations. This is referred in the rest of this paper as the conditional version of the test.

If n is too large to make the complete enumeration of the conditional distribution practical, the conditional P-value can still be approximated using a random subset of the 2^n possible sets. This method is used in a simulation study in the next section.

For large sample sizes, it is convenient to be able to rely on an asymptotic distribution. The rest of this section is devoted to this topic.

Let $F(x)$ be the cdf of θ_{12} under the null hypothesis. Introduce

$$\begin{aligned}
\sigma_{cc} &= E(s(Y_1+Y_2)s(Y_1+Y_3)\cos(\pi F(\theta_{12}))\cos(\pi F(\theta_{13}))) \\
\sigma_{ss} &= E(s(Y_1+Y_2)s(Y_1+Y_3)\sin(\pi F(\theta_{12}))\sin(\pi F(\theta_{13}))) \\
\sigma_{cs} &= E(s(Y_1+Y_2)s(Y_1+Y_3)\cos(\pi F(\theta_{12}))\sin(\pi F(\theta_{13})))
\end{aligned} \tag{5}$$

and define

$$\Sigma = \begin{pmatrix} \sigma_{cc} & \sigma_{cs} \\ \sigma_{cs} & \sigma_{ss} \end{pmatrix} \quad . \tag{6}$$

It is shown in the Appendix that Σ is, up to a multiplicative constant, the asymptotic covariance matrix of (A_n, B_n) under H_0. The following estimators of the above quantities will be needed. Let

$$\hat{\sigma}_{cc} = \frac{1}{\binom{n}{3}} \sum_{(i,j,k)\in I_3} s(Y_i+Y_j)s(Y_i+Y_k)\cos\left(\frac{\pi R_{ij}}{\binom{n}{2}}\right)\cos\left(\frac{\pi R_{ik}}{\binom{n}{2}}\right)$$

$$\hat{\sigma}_{ss} = \frac{1}{\binom{n}{3}} \sum_{(i,j,k)\in I_3} s(Y_i+Y_j)s(Y_i+Y_k)\sin\left(\frac{\pi R_{ij}}{\binom{n}{2}}\right)\sin\left(\frac{\pi R_{ik}}{\binom{n}{2}}\right)$$

$$\hat{\sigma}_{cs} = \frac{1}{2\binom{n}{3}} \sum_{(i,j,k)\in I_3} s(Y_i+Y_j)s(Y_i+Y_k)\left(\cos\left(\frac{\pi R_{ij}}{\binom{n}{2}}\right)\sin\left(\frac{\pi R_{ik}}{\binom{n}{2}}\right) + \cos\left(\frac{\pi R_{ik}}{\binom{n}{2}}\right)\sin\left(\frac{\pi R_{ij}}{\binom{n}{2}}\right)\right)$$

where I_3 denotes the set of unordered 3-tuples of distinct elements taken from the set $\{1,2,\ldots,n\}$; thus $\mathrm{Card}(I_3) = n(n-1)(n-2)/6$. Note that we set $\theta_{ij} = \theta_{ji}$ if $i > j$. Now, form the matrix $\hat{\Sigma}$ obtained from (6) by replacing the entries by their estimates. As will be seen in the Appendix, $\hat{\Sigma}$ is a consistent estimator of Σ. The test then consists in rejecting the hypothesis H_0 for large values of the statistic

$$S_n = (A_n, B_n)\hat{\Sigma}^{-1}(A_n, B_n)^T. \tag{7}$$

Once again, the test statistic is not distribution-free but has the following properties:

Theorem 2.1

a) S_n is affine-invariant.

b) Under H_0, under the assumption that Σ is nonsingular, that $(X_1 + X_2)/(Y_1 + Y_2)$ has a continuous distribution and $Pr(X_1 + X_2 = 0, Y_1 + Y_2 = 0) = 0$,

$$\frac{n}{4} S_n \xrightarrow{D} \chi^2_2 \text{ as } n \to \infty.$$

The test based on S_n performed using the asymptotic distribution is referred to as the asymptotic version of the test in the following. The proof of b) relies on the fact that under H_0, $\sqrt{n}|A_n^* - A_n| \xrightarrow{P} 0$ and $\sqrt{n}|B_n^* - B_n| \xrightarrow{P} 0$, where

$$A_n^* = \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} s(Y_j + Y_k) \cos(\pi F(\theta_{jk}))$$

$$B_n^* = \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} s(Y_j + Y_k) \sin(\pi F(\theta_{jk})). \tag{8}$$

We empirically found that the sample size required for the validity of the asymptotic distribution can be larger than usual. For example, with the standard bivariate normal distribution, a simulation was done using 5000 replications and a nominal level of 0.05. The following results were obtained, given in the form (sample size, observed level): (50, 0.018), (100, 0.0326), (150, 0.0402), (200, 0.0458), (250, 0.0464). In this case, it is seen that the test is maybe too conservative until n is approximately 150. Consequently, if the sample size is not very large, it is recommended to use the conditional version of the test. All computations in this paper were performed using Ox version 2.20; Doornik (1999). Programs that can compute S_n or the conditional P-value of the test based on C_n are available from the first author upon request. The Ox language is almost as fast as C so performing the conditional version of the test poses no problem, even for large sample sizes.

The methodology proposed in this paper is extended to the bivariate one-way layout and the simple regression models in Larocque *et al.* (2002). Bivariate sign tests based on the same projection idea are proposed and studied in Larocque, Tardif & van Eeden (2000).

3 Pitman efficiency and relationship with the work of Jan & Randles (1994)

In this section we will examine the relationship between our approach and the one of Jan & Randles (1994) (in the rest of this section, we will use JR to refer to this paper). First, the procedure of JR is applicable to any dimension, but we will focus on bivariate observations. Second, the assumption of elliptical symmetry is needed for the validity of the asymptotic null distribution of their statistic. Recall that we do not have to make such an assumption for the statistic S_n which needs only symmetry. However, the conditional test based on the JR statistics requires only symmetry to be valid.

For bivariate observations, the JR statistic is

$$SS = \frac{8}{n(n+1)^2} \sum_{i \le j} \sum_{k \le l} \cos(\pi \hat{p}_{ij,kl})$$

where $\hat{p}_{ij,kl} = (C_{ij,kl} + 2)/n \quad (0)$ if $(i,j) \neq (k,l)$ (if $(i,j) = (k,l)$) and where $C_{ij,kl}$ denotes the number of lines passing through the origin and one observation (excluding $\tilde{\mathbf{X}}_\mathbf{i}, \tilde{\mathbf{X}}_\mathbf{j}, \tilde{\mathbf{X}}_\mathbf{k}, \tilde{\mathbf{X}}_\mathbf{l}$) such that $\tilde{\mathbf{X}}_\mathbf{i} + \tilde{\mathbf{X}}_\mathbf{j}$ and $\tilde{\mathbf{X}}_\mathbf{k} + \tilde{\mathbf{X}}_\mathbf{l}$ are on opposite sides of the line formed.

Let

$$A_{JR} = \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} s(Y_j + Y_k) \cos\left(\frac{\pi R^*_{jk}}{n}\right)$$

$$B_{JR} = \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} s(Y_j + Y_k) \sin\left(\frac{\pi R^*_{jk}}{n}\right)$$

where R^*_{jk} is the rank of θ_{jk} among the set $\{\theta_{jk}, \theta_1, \ldots, \theta_n\}$ where $\theta_i = -\arctan(X_i/Y_i) + \pi/2$ is the angle ($\in [0, \pi)$) between the vector (X_i, Y_i) and the X-axis. Note that R^*_{jk} is an integer between 1 and $n+1$ and in consequence , the $n(n-1)/2$ R^*_{jk}s are not all distinct. It is shown in the Appendix that $SS/2$ is asymptotically equivalent to $n(A^2_{JR} + B^2_{JR})$ under the null hypothesis. Comparing (A_{JR}, B_{JR}) with the vector (A_n, B_n) (see (3)) on which the S_n test is based, we see that the only difference is in the way the θ_{jk}s are ranked. In our approach, the $n(n-1)/2$ angles θ_{jk} are ranked among themselves. On the other hand, the JR approach consists essentially (if n is large) in ranking the $n(n-1)/2$ angles among the n θ_is.

Now, a similar argument to the one used in the proof of Lemma A.5 shows that, under H_0, the vectors (A_{JR}, B_{JR}) and (A^*_{JR}, B^*_{JR}) are asymptotically equivalent, that is the difference between the two vectors converges to (0,0) in probability, where

$$A^*_{JR} = \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} s(Y_j + Y_k) \cos(\pi F_1(\theta_{jk}))$$

$$B^*_{JR} = \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} s(Y_j + Y_k) \sin(\pi F_1(\theta_{jk}))$$

with $F_1(x)$ being the cdf of θ_1 under the null hypothesis. Let us compare (A^*_{JR}, B^*_{JR}) with (A^*_n, B^*_n) given by (8). The only difference between those vectors is the fact that the former uses F_1, the null cdf of θ_1, while the latter uses F, the null cdf of θ_{12}. In general, these two cdf are not equal. This suggests that the approach used in this paper and that used in JR are not generally equivalent. On the other hand, if the observations come from a spherical distribution, we have $F(x) = F_1(x) = x/\pi, \quad 0 < x < \pi$. That is, both θ_1 and θ_{12} are uniformly distributed on $[0, \pi]$ under H_0. Indeed, for F, this follows from the fact that $\tilde{\mathbf{X}}_\mathbf{1} + \tilde{\mathbf{X}}_\mathbf{2}$ is spherically symmetric if $\tilde{\mathbf{X}}_\mathbf{1}$ and $\tilde{\mathbf{X}}_\mathbf{2}$ are both i.i.d. spherically symmetric. This means that the JR approach and ours are asymptotically equivalent under H_0 as well as under a contiguous sequence of alternatives for spherical distributions. But this also means that they are asymptotically equivalent for elliptical observations since the two approaches are affine-invariant. Consequently, the two approaches are equivalent in terms of Pitman efficiency for elliptical distributions. Furthermore, S_n and SS are asymptotically equivalent to the spatial signed-rank test for spherical distributions; see Möttönen & Oja (1995) and Möttönen, Oja & Tienari (1997). The asymptotic Pitman efficiency of the S_n test relative to Hotelling's test is thus given in Jan & Randles (1994 Table 5.1) for the power family of elliptical distributions and in Möttönen, Oja & Tienari (1997 Table 1) for the bivariate t distribution. The relevant entries of those tables are reproduced in Table 1 for convenience. These ARE values show that the test based on S_n performs very well when compared with Hotelling's T^2 test. In particular, it outperforms the T^2 test for heavy-tailed distributions and remains quite competitive for distributions with lighter tails. The ARE for the normal distribution (where the T^2 test is optimal) is 0.967. In comparison, the ARE of the Wilcoxon signed-rank test relative to the t test is 0.955 for the normal distribution in the univariate

Table 1: Asymptotic efficiency of the test S_n relative to Hotelling's test for the bivariate t and exponential power distributions

Bivariate t distribution	Degrees of freedom						
	3	4	6	8	10	20	∞
ARE(S_n,Hotelling)	1.953	1.435	1.187	1.108	1.071	1.011	0.967
Exponential power distribution	ν						
	1/3	1/2	2/3	1	3/2	2	3
ARE(S_n,Hotelling)	1.860	1.296	1.106	0.967	0.904	0.883	0.871

case.

To conclude this section, the fact that the JR approach and ours are not equivalent for non-elliptical distributions is illustrated with a small simulation study. The conditional versions of both tests are used, that is, we compare the tests based on $A_n^2 + B_n^2$ and $A_{JR}^2 + B_{JR}^2$. Note that $A_{JR}^2 + B_{JR}^2$ is not exactly the conditional test proposed in JR. We use a slightly modified version here because it takes less time to compute. The computing costs of performing numerous conditional tests with the original version of the test would have been prohibitive. It also allows us to focus directly on the two different ways of ranking the angles. The results of Hotelling's test are also reported for comparison. The sample size is $n = 20$. Three distributions are used. The first is the standard bivariate normal distribution. Since this is an elliptical distribution, we expect the power of the two signed-rank tests to be close. The other distribution is non-elliptical. It is simply the distribution of (X, Y) where X and Y are independent, X being uniformly distributed over $(-1, 1)$ and Y having the standard Cauchy distribution. We call this distribution uniform-Cauchy. The null hypothesis (amount of shift (0,0)) and three sets of shifts are considered for the bivariate normal distribution. They take the form (c, c) for different values of c. The null hypothesis and six sets of shifts are considered for the uniform-Cauchy distribution. Three shifts take the form $(c, 0)$, i.e. they lie in the uniform direction and the other three take the form $(0, c)$, i.e. they lie in the

Table 2: Observed probability of rejecting H_0 (n=20, 1000 replications)

Population	Amount of shift	Statistics		
		Hotelling's	Jan and Randles	$A_n^2 + B_n^2$
Bivariate normal	(0,0)	0.043	0.052	0.049
	(0.25,0.25)	0.216	0.202	0.209
	(0.4,0.4)	0.523	0.495	0.516
	(0.6,0.6)	0.892	0.874	0.877
Uniform-Cauchy	(0,0)	0.038	0.047	0.044
	(0.18,0)	0.188	0.135	0.171
	(0.33,0)	0.522	0.369	0.467
	(0.42,0)	0.767	0.603	0.704
	(0,0.5)	0.064	0.150	0.131
	(0,1.8)	0.329	0.659	0.540
	(0,2.8)	0.508	0.818	0.709

Cauchy direction. Each power is estimated with 1,000 replications. Each time a conditional test is performed, its P-value is approximated with 1,000 random sets $\tilde{\mathbf{X}}_{\mathbf{\Gamma}}$.

The results are reported in Table 2. As expected, for the bivariate normal distribution, the power of two signed-rank tests are very close. Moreover, they are only slightly less powerful than Hotelling's test. This is in accordance with the fact that the ARE of (the asymptotic versions of) the signed-rank tests relative to Hotelling's test is 0.967. The fact that the two signed-rank tests are not equivalent for non-elliptical distributions is clearly seen with the uniform-Cauchy distribution. The test based on $A_n^2 + B_n^2$ is the more powerful of the two when the shifts take the form $(c, 0)$ while the opposite is true for shifts of the form $(0, c)$.

4 Example and concluding remarks

As an example, we use the immune system data set analyzed in Randles (2000). The data comprises CD45RA T and CD45RO T cell counts of 36 HIV-positive newborn infants measured at birth and after 24 weeks of ritonavir therapy and are taken from Sleasman *et al.* (1999). The same tests used in the simulation study of the preceding section, that is the conditional versions of the tests based on $A_n^2 + B_n^2$ and $A_{JR}^2 + B_{JR}^2$ (but with 100000 random sets this time) and Hotelling's test, were applied to the changes in CD45RA and CD45RO given in Table 3. The results are reported in Table 4. All tests indicate strong significance. To examine the stability of the results, the same tests were applied to a modified data set using the same alteration as in Randles (2000), i.e. changing the first vector of observation from (242, 1708) to (–1758, –4292). We see that the two signed-rank tests remain significant, but that Hotelling's test is no longer significant.

In summary, we have proposed two different ways to perform a signed-rank test for the bivariate location problem. The two statistics are affine-invariant and the tests performs well compared to Hotelling's test even under normality. Furthermore, no additional assumption (such as elliptical symmetry of the observations) is needed to use the asymptotic version of the test. The relationship with the approach of Jan & Randles (1994) has been noted and examined. It is seen that the two approaches are equivalent in the Pitman ARE sense, but only for elliptical distributions. Unlike the asymptotic version of the Jan & Randles (1994) test, the asymptotic version of the new test does not require the assumption of elliptical symmetry as only symmetry is needed. The new test is also asymptotically equivalent to the spatial signed-rank test of Möttönen & Oja (1995) for spherical distributions. But the latter is only rotation-invariant (not scale-invariant) and may be affected by a large difference in scale between the two coordinates.

Acknowledgement

This research was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada and by the Fonds pour la formation de chercheurs et l'aide à la recherche (FCAR) of the Province of Québec.

Table 3: Immune system responses data

Patient	Change in CD45RA	Change in CD45RO	Patient	Change in CD45RA	Change in CD45RO
1	242	1708	19	–9	36
2	569	569	20	158	234
3	270	757	21	76	122
4	–25	499	22	15	24
5	309	231	23	3	36
6	22	338	24	93	71
7	–42	26	25	160	44
8	–233	119	26	66	128
9	206	163	27	180	155
10	–106	–186	28	237	85
11	55	54	29	105	76
12	85	48	30	16	6
13	30	50	31	167	364
14	194	525	32	–10	–18
15	–87	–110	33	–61	–21
16	159	148	34	–7	–2
17	29	102	35	15	32
18	89	364	36	160	188

Table 4: Results for the immune system responses data

	Statistic		
	Hotelling	Jan and Randles	$A_n^2 + B_n^2$
P-value (original data)	0.0011	0.0002	0.0002
P-value (altered data)	0.5546	0.0017	0.0039

A APPENDIX

Proof of Theorem 2.1 a):

It is sufficient to show that S_n is scale-invariant, rotation-invariant and invariant for a reflection about the X-axis since any linear transformation (by a nonsingular 2×2 matrix) can be decomposed into a sequence of those types of transformations. This will be done successively in three lemmas.

Note: in the three lemmas below, any random variable with the (t) symbol represents the transformed random variable.

Lemma A.1 *$\hat{\sigma}_{cc}$, $\hat{\sigma}_{ss}$, $\hat{\sigma}_{cs}$, A_n, B_n, and S_n are scale-invariant.*

Proof: The transformed points are

$$(X_i^{(t)}, Y_i^{(t)}) = (aX_i, bY_i) \quad , \quad i = 1, \ldots, n,$$

where $a > 0$ and $b > 0$ are two constants. The signs $s(Y_j + Y_k)$ are not affected by such a transformation. On the other hand,

$$\theta_{jk}^{(t)} = -\arctan\left(\frac{aX_j + aX_k}{bY_j + bY_k}\right) + \frac{\pi}{2} = -\arctan\left(\frac{a}{b}\left(\frac{X_j + X_k}{Y_j + Y_k}\right)\right) + \frac{\pi}{2}.$$

Thus, $R_{jk}^{(t)}$ (the rank of $\theta_{jk}^{(t)}$ among the $n(n-1)/2$ $\theta_{jk}^{(t)}$s) equals R_{jk} since arctan is a strictly increasing function. Since $\hat{\sigma}_{cc}$, $\hat{\sigma}_{ss}$, $\hat{\sigma}_{cs}$, A_n and B_n are only functions of the $s(Y_j + Y_k)$s and

R_{jk}s, they are scale-invariant. Also, since S_n is a function of only those random variables, it is scale-invariant which concludes the proof.○

Now, consider a reflection of the points about the X-axis. The transformed observations are

$$(X_i^{(t)}, Y_i^{(t)}) = (X_i, -Y_i) \quad , \quad i = 1, \ldots, n.$$

Let $c = \cos(2\pi/(n(n-1))$ and $s = \sin(2\pi/(n(n-1))$.

Lemma A.2 *If the observations are reflected about the X-axis, then*

(i) $A_n^{(t)} = cA_n + sB_n$, $B_n^{(t)} = -cB_n + sA_n$, $\hat{\sigma}_{cc}^{(t)} = c^2\hat{\sigma}_{cc} + 2cs\hat{\sigma}_{cs} + s^2\hat{\sigma}_{ss}$,
$\hat{\sigma}_{ss}^{(t)} = c^2\hat{\sigma}_{ss} - 2cs\hat{\sigma}_{cs} + s^2\hat{\sigma}_{cc}$, $\hat{\sigma}_{cs}^{(t)} = -c^2\hat{\sigma}_{cs} + cs\hat{\sigma}_{cc} - cs\hat{\sigma}_{ss} + s^2\hat{\sigma}_{cs}$ $|\hat{\Sigma}^{(t)}| = |\hat{\Sigma}|$.

(ii) $S_n^{(t)} = S_n$ *(i.e.* S_n *is invariant for reflections about the X-axis).*

Proof: On one hand, we have

$$s(Y_j^{(t)} + Y_k^{(t)}) = -s(Y_j + Y_k) \quad \forall \quad j, k. \tag{9}$$

On the other hand,

$$\theta_{jk}^{(t)} = -\arctan\left(\frac{X_j + X_k}{-(Y_j + Y_k)}\right) + \frac{\pi}{2} = \arctan\left(\frac{X_j + X_k}{Y_j + Y_k}\right) + \frac{\pi}{2} = -\theta_{jk} + \pi.$$

Thus,

$$R_{jk}^{(t)} = \binom{n}{2} + 1 - R_{jk}. \tag{10}$$

Using (9), (10) and elementary trigonometric identities, we have

$$\cos\left(\frac{\pi R_{jk}^{(t)}}{\binom{n}{2}}\right) = -c\cos\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right) - s\sin\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right) \tag{11}$$

and

$$\sin\left(\frac{\pi R_{jk}^{(t)}}{\binom{n}{2}}\right) = c\sin\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right) - s\cos\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right). \tag{12}$$

Part (i) is shown with straightforward algebra by using (11) and (12). As for (ii), it can be shown, again with long but straightforward algebra, that

$$S_n^{(t)} = \frac{1}{|\hat{\Sigma}^{(t)}|}(A_n^{(t)}, B_N^{(t)}) \begin{pmatrix} \hat{\sigma}_{ss}^{(t)} & -\hat{\sigma}_{cs}^{(t)} \\ -\hat{\sigma}_{cs}^{(t)} & \hat{\sigma}_{cc}^{(t)} \end{pmatrix} \begin{pmatrix} A_n^{(t)} \\ B_n^{(t)} \end{pmatrix} = S_n$$

which completes the proof of Lemma A.2.○

Lemma A.3 *S_n is rotation-invariant.*

Proof: Suppose that the rotation is such that one and only one of the $n(n-1)/2$ points $(X_j + X_k, Y_j + Y_k)$ will cross the X-axis and suppose wlog that $(X_1 + X_2, Y_1 + Y_2)$ is the point in question. This point has the smallest angle θ_{12}. After the rotation, the point $(X_1^{(t)} + X_2^{(t)}, Y_1^{(t)} + Y_2^{(t)})$ now has the greatest angle (i.e. $R_{12}^{(t)} = n(n-1)/2$) and all the other points lose a rank (i.e. $R_{jk}^{(t)} = R_{jk} - 1 \quad \forall (j,k) \neq (1,2)$). Furthermore, $s(Y_1^{(t)} + Y_2^{(t)}) = -s(Y_1 + Y_2)$ as the point crossed the X-axis and $s(Y_j^{(t)} + Y_k^{(t)}) = s(Y_j + Y_k)\ \forall (j,k) \neq (1,2)$ as all these points did not cross the X-axis. Using these facts, an argument similar to the one in the proof of Lemma A.2 shows that $S_n^{(t)} = S_n$ in that case. The rest of the proof is clear: if the rotation is such that exactly two of the points will cross the X-axis, then the value of S_n will remain unchanged and so on. This concludes the proof of Lemma A.3 and of Theorem 2.1 a).○

Proof of Theorem 2.1 b):

Let us begin with a first preliminary result.

Lemma A.4

(a) $s(Y_1 + Y_2)$ *and* θ_{12} *are independent.*

(b)

$$\lim_{n\to\infty} \max_{1\leq j\neq k\leq n} E\left(\frac{R_{jk}}{\binom{n}{2}} - F(\theta_{jk})\right)^2 = 0$$

(c)

$$\lim_{n\to\infty} n \max_{\substack{j,k,l,m \\ \text{all distinct}}} \Big| E\Big(s(Y_j + Y_k)s(Y_l + Y_m)$$

$$\left(\cos\left(\frac{\pi}{\binom{n}{2}} R_{jk}\right) - \cos(\pi F(\theta_{jk}))\right)\left(\cos\left(\frac{\pi}{\binom{n}{2}} R_{lm}\right) - \cos(\pi F(\theta_{lm}))\right)\Big)\Big| = 0.$$

Note that part (c) above remains valid if we replace the cos by sin.

Proof:

(a) Since $(X_1 + X_2, Y_1 + Y_2)$ and $-(X_1 + X_2, Y_1 + Y_2)$ are identically distributed, it is easy to show that $(X_1 + X_2)/(Y_1 + Y_2)$ and $s(Y_1 + Y_2)$ are independent. The fact that θ_{12} is a function of $(X_1 + X_2)/(Y_1 + Y_2)$ shows part (a).

(b) Write

$$R_{jk} = \sum_{(l,m)\in H} \varphi(\theta_{jk} > \theta_{lm}) \quad + 1$$

where φ is the indicator function and $H = \{(l,m) : \quad 1 \le l < m \le n; \quad (l,m) \ne (j,k)\}$. Now

$$E(R_{jk}|\theta_{jk}) = \sum_{(l,m)\in H} E(\varphi(\theta_{jk} > \theta_{lm})|\theta_{jk}) \quad + 1 = \binom{n-2}{2} F(\theta_{jk}) + O(n).$$

Similarly, we have

$$E(R_{jk}^2|\theta_{jk}) = \binom{n-2}{2}\binom{n-4}{2} F^2(\theta_{jk}) + O(n^3).$$

Thus

$$\begin{aligned} &\left| E\left(\left(\frac{R_{jk}}{\binom{n}{2}} - F(\theta_{jk}) \right)^2 \Big| \theta_{jk} \right) \right| \\ &= \left| \frac{\binom{n-2}{2}\binom{n-4}{2}}{\binom{n}{2}^2} F^2(\theta_{jk}) - 2\frac{\binom{n-2}{2}}{\binom{n}{2}} F^2(\theta_{jk}) + F^2(\theta_{jk}) + O(1/n) \right| \\ &= |o(1)F^2(\theta_{jk}) + O(1/n)| = o(1) \quad \text{independently of} \quad j,k. \end{aligned}$$

This proves part (b).

(c) For j,k,l,m distinct integers $(j < k, l < m)$, define

$$R_{jk}^{(lm)} = \sum_{(r,t)\in I_4} \varphi(\theta_{jk} > \theta_{rt})$$

where $I_4 = \{(r,t) : \quad 1 \le r < t \le n; \quad r \ne j,k,l,m; \quad t \ne j,k,l,m\}$. It is easy to see that

$$|R_{jk} - R_{jk}^{(lm)}| = O(n) \quad \text{independently of} \quad j,k,l,m. \tag{13}$$

From this and part (b) above it can be easily seen that

$$\lim_{n\to\infty} \max_{\substack{j,k,l,m \\ \text{all distinct}}} E\left(\left(\frac{R_{jk}^{(lm)}}{\binom{n}{2}} - F(\theta_{jk}) \right)^2 \right) = 0. \tag{14}$$

Write

$$\cos\left(\frac{\pi}{\binom{n}{2}}R_{jk}\right) = \cos\left(\frac{\pi}{\binom{n}{2}}R_{jk}^{(lm)} + \xi_{jk}^{(lm)}\right)$$

where $\xi_{jk}^{(lm)} = 2\pi(R_{jk} - R_{jk}^{(lm)}/(n(n-1))$ satisfies $|\xi_{jk}^{(lm)}| = O(1/n)$ from (13) above. A first order Taylor series expansion of the last term gives

$$\cos\left(\frac{\pi}{\binom{n}{2}}R_{jk}^{(lm)}\right) - \sin\left(\frac{\pi}{\binom{n}{2}}R_{jk}^{(lm)}\right)\xi_{jk}^{(lm)*}$$

where

$$|\xi_{jk}^{(lm)*}| \leq |\xi_{jk}^{(lm)}| = O(1/n). \tag{15}$$

Introduce the notation Max$(\cdot)$ for the maximum of the argument over the indices j,k,l,m such that j,k,l,m are all distinct. Using the above and expanding the terms we can write

$$\begin{aligned}
& n\,\mathrm{Max}\Big|E\Big(s(Y_j+Y_k)s(Y_l+Y_m) \\
& \qquad \left(\cos\left(\frac{\pi}{\binom{n}{2}}R_{jk}\right) - \cos(\pi F(\theta_{jk}))\right)\left(\cos\left(\frac{\pi}{\binom{n}{2}}R_{lm}\right) - \cos(\pi F(\theta_{lm}))\right)\Big)\Big| \\
\leq\ & n\mathrm{Max}\Big|E\Big(\left(\cos\left(\frac{\pi}{\binom{n}{2}}R_{jk}^{(lm)}\right) - \cos(\pi F(\theta_{jk}))\right)\left(\cos\left(\frac{\pi}{\binom{n}{2}}R_{lm}^{(jk)}\right) - \cos(\pi F(\theta_{lm}))\right)\times \\
& E[s(Y_j+Y_k)S(Y_l+Y_m)|\theta_{jk},\theta_{lm},\ \text{all } \theta_{rt}\text{s except those indexed by } j,k,l,m]\Big)\Big| \qquad (16) \\
& +n\,\mathrm{Max}\left|E\left(n\xi_{jk}^{(lm)*}\sin\left(\frac{\pi}{\binom{n}{2}}R_{jk}^{(lm)}\right)\left(\cos\left(\frac{\pi}{\binom{n}{2}}R_{lm}^{(jk)}\right) - \cos(\pi F(\theta_{lm}))\right)\right)\right| \\
& +n\,\mathrm{Max}\left|E\left(n\xi_{lm}^{(jk)*}\sin\left(\frac{\pi}{\binom{n}{2}}R_{lm}^{(jk)}\right)\left(\cos\left(\frac{\pi}{\binom{n}{2}}R_{jk}^{(lm)}\right) - \cos(\pi F(\theta_{jk}))\right)\right)\right| \\
& +n\,\mathrm{Max}\left|E\left(n\xi_{lm}^{(jk)*}\xi_{jk}^{(lm)*}\sin\left(\frac{\pi}{\binom{n}{2}}R_{lm}^{(jk)}\right)\sin\left(\frac{\pi}{\binom{n}{2}}R_{jk}^{(lm)}\right)\right)\right|.
\end{aligned}$$

For the first term on the right-hand side of (16), we have

$$\begin{aligned}
& E(s(Y_j+Y_k)s(Y_l+Y_m)|\theta_{jk},\theta_{lm},\ \text{all the } \theta_{rt}\text{s except those indexed by } j,k,l,m) \\
=\ & E(s(Y_j+Y_k)s(Y_l+Y_m)|\theta_{jk},\theta_{lm}) \\
=\ & E(s(Y_j+Y_k)|\theta_{jk})E(s(Y_l+Y_m)|\theta_{lm}) \quad \text{as } j,k,l,m \text{ are distinct} \\
=\ & E(s(Y_j+Y_k))E(s(Y_l+Y_m)) = 0.
\end{aligned}$$

Where the last equality follows from the independence between θ_{jk} and $s(Y_j + Y_k)$. Using the Cauchy-Schwarz inequality, the second term on the right-hand side of (16) is bounded above by

$$\begin{aligned} &\mathrm{Max}\Big|\Big\{E\left(n^2(\xi_{jk}^{(lm)*})^2\sin^2\left(\frac{\pi}{\binom{n}{2}}R_{jk}^{(lm)}\right)\right)\times \\ &\qquad E\left(\left(\cos\left(\frac{\pi}{\binom{n}{2}}R_{lm}^{(jk)}\right)-\cos(\pi F(\theta_{lm}))\right)^2\right)\Big\}^{1/2}\Big| \qquad (17) \\ \leq\ & O(1)\max_{\substack{j,k,l,m\\ \text{all distinct}}}\left\{E\left(\left(\cos\left(\frac{\pi}{\binom{n}{2}}R_{lm}^{(jk)}\right)-\cos(\pi F(\theta_{lm}))\right)^2\right)\right\}^{1/2} \end{aligned}$$

as $n^2(\xi_{jk}^{(lm)*})^2\sin^2(2\pi R_{jk}^{(lm)}/(n(n-1))) = O(1)$ independently of j,k,l,m; see (15). The term (17) then converges to 0 as $n\to\infty$ from (14). The last two terms on the right-hand side of (16) are treated similarly and this concludes the proof of part (c).◯

Lemma A.5

$$\sqrt{n}|A_n^* - A_n| \xrightarrow{P} 0 \quad \textit{and} \quad \sqrt{n}|B_n^* - B_n| \xrightarrow{P} 0.$$

Proof: The first part will be proven by showing that

$$\lim_{n\to\infty} nE(A_n^* - A_n)^2 = 0.$$

First, define the sets $I_1 = \{j,k,l,m = 1,2,\ldots,n:\ j<k;\ l<m;\ j,k,l,m \quad \text{all distinct}\}$ and $I_2 = \{j,k,l,m = 1,2,\ldots,n:\ j<k;\ l<m;\ j,k,l,m \quad \text{not all distinct}\}$. Note that

$$\begin{aligned} nE(A_n^* - A_n)^2 \ =\ & \frac{4}{n(n-1)^2}\sum_{j,k,l,m\in I_1}E(s(Y_j+Y_k)s(Y_l+Y_m)\times \\ & (\cos\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right)-\cos(\pi F(\theta_{jk})))(\cos\left(\frac{\pi R_{lm}}{\binom{n}{2}}\right)-\cos(\pi F(\theta_{lm})))) \\ +\ & \frac{4}{n(n-1)^2}\sum_{j,k,l,m\in I_2}E(s(Y_j+Y_k)s(Y_l+Y_m)\times \\ & (\cos\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right)-\cos(\pi(\theta_{jk})))(\cos\left(\frac{\pi R_{lm}}{\binom{n}{2}}\right)-\cos(\pi F(\theta_{lm})))) \end{aligned} \qquad (18)$$

Now, the first term on the right-hand side of (18) is bounded above by

$$4n\max_{j,k,l,m\in I_1}\left|E(s(Y_j+Y_k)s(Y_l+Y_m)(\cos\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right)-\cos(\pi F(\theta_{jk})))(\cos\left(\frac{\pi R_{lm}}{\binom{n}{2}}\right)-\cos(\pi F(\theta_{lm}))))\right|$$

which tends to 0 by lemma A.4.c . Using the Cauchy-Schwarz inequality, the second term on the right-hand side of (18) is bounded above by

$$4 \max_{j,k,l,m \in I_2} \left\{ E\left(\cos\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right) - \cos(\pi F(\theta_{jk}))\right)^2 E\left(\cos\left(\frac{\pi R_{lm}}{\binom{n}{2}}\right) - \cos(\pi F(\theta_{lm}))\right)^2 \right\}^{1/2}$$

$$\leq 4 \max_{1 \leq j \neq k \leq n} E\left(\cos\left(\frac{\pi R_{jk}}{\binom{n}{2}}\right) - \cos(\pi F(\theta_{jk}))\right)^2$$

$$\leq 4K\pi^2 \max_{1 \leq j \neq k \leq n} E\left(\frac{R_{jk}}{\binom{n}{2}} - F(\theta_{jk}))\right)^2$$

as $|\cos(a) - \cos(b)| \leq K|a - b|$ for a certain $K > 0$. Since this last term converges to 0 by lemma A.4.b , this completes the proof of the first part of the theorem. The second part can be proven in a similar way. ○

We can now proceed with the final part of the proof of Theorem 2.1 b). Note that both A_n^* and B_n^* are U-statistics with respective kernels

$$h_A((x_1, y_1), (x_2, y_2)) = s(y_1 + y_2) \cos\left(\pi F\left(-\arctan\left(\frac{x_1 + x_2}{y_1 + y_2}\right) + \frac{\pi}{2}\right)\right)$$

and

$$h_B((x_1, y_1), (x_2, y_2)) = s(y_1 + y_2) \sin\left(\pi F\left(-\arctan\left(\frac{x_1 + x_2}{y_1 + y_2}\right) + \frac{\pi}{2}\right)\right).$$

Since $E(s(Y_1 + Y_2)) = 0$ under the null hypothesis, we have

$$E(A_n^*) = E(s(Y_1 + Y_2)\cos(\pi F(\theta_{12}))) = E(s(Y_1 + Y_2))E(\cos(\pi F(\theta_{12}))) = 0$$

from the independence between $s(Y_1 + Y_2)$ and θ_{12}. Also, $E(B_n^*) = 0$ under H_0. Suppose that Σ is nonsingular. Standard results on the theory of U-statistics then show that, under the null hypothesis,

$$\sqrt{n}(A_n^*, B_n^*)^T \xrightarrow{D} N(\mathbf{0}, 4\Sigma) \quad \text{as} \quad n \to \infty; \tag{19}$$

In view of (19) and the last lemma, we only need to show that $\tilde{\Sigma}$ is a consistent estimator of Σ. In fact, it is sufficient to show that

$$|\hat{\sigma}_{cc} - \sigma_{cc}| \xrightarrow{P} 0, \quad |\hat{\sigma}_{ss} - \sigma_{ss}| \xrightarrow{P} 0, \quad \text{and} \quad |\hat{\sigma}_{cs} - \sigma_{cs}| \xrightarrow{P} 0$$

We only show the first result as the other two can be treated by using the same kind of argument. First, introduce

$$\tilde{\sigma}_{cc} = \frac{1}{\binom{n}{3}} \sum_{(i,j,k)\in I_3} s(Y_i + Y_j)s(Y_i + Y_k)\cos(\pi F(\theta_{ij}))\cos(\pi F(\theta_{ik})).$$

Obviously,

$$|\hat{\sigma}_{cc} - \sigma_{cc}| \leq |\hat{\sigma}_{cc} - \tilde{\sigma}_{cc}| + |\tilde{\sigma}_{cc} - \sigma_{cc}|. \tag{20}$$

Considerations similar to those used in the proof of the last lemma lead to $|\hat{\sigma}_{cc} - \tilde{\sigma}_{cc}| \xrightarrow{P} 0$. As for the second term on the right-hand side of (20), note that we can write

$$\begin{aligned}
\tilde{\sigma}_{cc} = \; & \tfrac{1}{\binom{n}{3}} \textstyle\sum_{(i,j,k)\in I_3} \tfrac{1}{3} \quad (s(Y_i + Y_j)s(Y_i + Y_k)\cos(\pi F(\theta_{ij}))\cos(\pi F(\theta_{ik})) \\
& \qquad + s(Y_i + Y_k)s(Y_j + Y_k)\cos(\pi F(\theta_{ik}))\cos(\pi F(\theta_{jk})) \\
& \qquad + s(Y_i + Y_j)s(Y_j + Y_k)\cos(\pi F(\theta_{ij}))\cos(\pi F(\theta_{jk}))) \\
= \; & \tfrac{1}{\binom{n}{3}} \textstyle\sum_{(i,j,k)\in I_3} \quad h((X_i, Y_i), (X_j, Y_j), (X_k, Y_k)),
\end{aligned}$$

say, which is a U-statistic of order 3 with kernel h satisfying $E(|h|) < \infty$ since $|h|$ is bounded. Hence $\tilde{\sigma}_{cc} \xrightarrow{P} E(h) = (\sigma_{cc} + \sigma_{cc} + \sigma_{cc})/3 = \sigma_{cc}$, which completes the proof.◯

Asymptotic null equivalence between $SS/2$ and $n(A_{JR}^2 + B_{JR}^2)$ (Section 3):

Adopt the convention that $\theta_{ii} = \theta_i$. In order to simplify the notation in what follows, define $s_{ij} = s(Y_i + Y_j)$ and $R_{ij}^{(ijkl)}$ as the rank of θ_{ij} among $\theta_1, \ldots, \theta_n$ (except $\theta_i, \theta_j, \theta_k, \theta_l$) and θ_{ij}, θ_{kl} $(i \leq j\,,\, k \leq l\,,\, (i,j) \neq (k,l))$. Note that if the four indices are distinct, then $1 \leq R_{ij}^{(ijkl)} \leq n-2$, if three are distinct, then $1 \leq R_{ij}^{(ijkl)} \leq n-1$ and if only two are distinct, then $1 \leq R_{ij}^{(ijkl)} \leq n$.

First, suppose that i, j, k, l are distinct. Then

$$C_{ij,kl} = \begin{cases} |R_{ij}^{(ijkl)} - R_{kl}^{(ijkl)}| - 1 & \text{if} \quad s_{ij}s_{kl} = 1 \\ n - |R_{ij}^{(ijkl)} - R_{kl}^{(ijkl)}| - 1 & \text{if} \quad s_{ij}s_{kl} = -1. \end{cases}$$

This in turn gives

$$\hat{p}_{ij,kl} = \begin{cases} \frac{|R_{ij}^{(ijkl)} - R_{kl}^{(ijkl)}| + 1}{n} & \text{if} \quad s_{ij}s_{kl} = 1 \\ 1 - \frac{|R_{ij}^{(ijkl)} - R_{kl}^{(ijkl)}| - 1}{n} & \text{if} \quad s_{ij}s_{kl} = -1. \end{cases}$$

A two-term Taylor series expansion then gives

$$\begin{aligned}\cos(\pi\hat{p}_{ij,kl}) &= s_{ij}s_{kl}\cos\left(\frac{\pi}{n}s_{ij}s_{kl}+\frac{\pi}{n}|R_{ij}^{(ijkl)}-R_{kl}^{(ijkl)}|\right)\\ &= s_{ij}s_{kl}\cos\left(\frac{\pi}{n}|R_{ij}^{(ijkl)}-R_{kl}^{(ijkl)}|\right)-\frac{\pi}{n}\sin\left(\frac{\pi}{n}|R_{ij}^{(ijkl)}-R_{kl}^{(ijkl)}|\right)+O(n^2).\end{aligned}$$

Similarly, we have

$$\cos(\pi\hat{p}_{ij,kl}) = s_{ij}s_{kl}\cos\left(\frac{\pi}{n}|R_{ij}^{(ijkl)}-R_{kl}^{(ijkl)}|\right)+O(n)$$

if three or less indices from i,j,k,l are distinct. Thus

$$\begin{aligned}\frac{SS}{2} &= \frac{4}{n(n+1)^2}\sum_{i\le j}\sum_{k\le l}\cos(\pi\hat{p}_{ij,kl})\\ &= \frac{4}{n(n+1)^2}\sum_{i\le j}\sum_{k\le l}s(Y_i+Y_j)s(Y_k+Y_l)\cos\left(\frac{\pi}{n}|R_{ij}^{(ijkl)}-R_{kl}^{(ijkl)}|\right) \qquad (21)\\ &\quad -\frac{4\pi}{n^2(n+1)^2}\sum_{\substack{j,k,l,m\\ \text{all distinct}}}\sin\left(\frac{\pi}{n}|R_{ij}^{(ijkl)}-R_{kl}^{(ijkl)}|\right)+O(1/n).\end{aligned}$$

Now, the second term on the right-hand side of (21) converges in probability, under H_0, to $-4\pi E_{H_0}(\sin(\pi(F_1(\theta_{12})-F_1(\theta_{34}))))=0$. Thus, $SS/2$ and

$$\frac{n}{\binom{n}{2}^2}\sum_{i\le j}\sum_{k\le l}s(Y_i+Y_j)s(Y_k+Y_l)\cos\left(\frac{\pi}{n}(R_{ij}^{(ijkl)}-R_{kl}^{(ijkl)})\right) \qquad (22)$$

are asymptotically equivalent under the null hypothesis. Now, under H_0, we can neglect the terms $i=j$ and $k=l$ in the summations above. This, along with the fact that $|R_{ij}^{*}-R_{ij}^{(ijkl)}|\le 6$ easily entails that (22) is also asymptotically equivalent, under H_0, to

$$\frac{n}{\binom{n}{2}^2}\sum_{i<j}\sum_{k<l}s(Y_i+Y_j)s(Y_k+Y_l)\cos\left(\frac{\pi}{n}(R_{ij}^{*}-R_{kl}^{*})\right)=n(A_{JR}^2+B_{JR}^2)$$

which completes the proof.○

References

Bennett, B. M. (1964). A Bivariate Signed-Rank Test. *J. Roy. Statist. Soc. Ser.* B, **26**, 457-461.

Brown, B. M. & Hettmansperger, T. P. (1987). Affine Invariant Rank Methods in the Bivariate Location Model. *J. Roy. Statist. Soc. Ser.* B, **49**, 301-310.

Doornik, J. A. (1999). *Object-Oriented Matrix Programming Using Ox*, 3rd ed. Timberlake Consultants Press and Oxford, London.

Hettmansperger, T. P. & McKean, J. W. (1998). *Robust Nonparametric Statistical Methods*. Arnold. London.

Hössjer, O. & Croux, C. (1995). Generalizing Univariate Signed Rank Statistics for Testing and Estimating a Multivariate Location Parameter. *J. Nonparametr. Statist.*, **4**, 293-308.

Jan, S.-L. & Randles, R. H. (1994). A Multivariate Signed Sum Test for the One-Sample Location Problem. *J. Nonparametr. Statist.*, **4**, 49-63.

Larocque, D., Tardif, S. & van Eeden, C. (2000). Bivariate Sign Tests Based on the Sup, L_1 and L_2 Norms. *Ann. Inst. Statist. Math.* **52**, 488-506.

Larocque, D., Tardif, S. & van Eeden, C. (2002). Affine-Invariant Rank Tests for the Bivariate One-Way Layout and for Simple Regression Models. To appear in *Statistics*.

Möttönen, J. & Oja, H. (1995). Multivariate Spatial Sign and Rank Methods. *J. Nonparametr. Statist.*, **5**, 201-213.

Möttönen, J., Oja, H. & Tienari, J. (1997). On The Efficiency of Multivariate Spatial Sign and Rank Tests. *Ann. Statist.*, **25**, 542-552.

Peters, D. & Randles, R. H. (1990). A Multivariate Signed-Rank Test for the One-Sample Location Problem. *J. Amer. Statist. Assoc.*, **85**, 552-557.

Randles, R. H. (2000). A Simpler, Affine-Invariant, Multivariate, Distribution-Free Sign Test. *J. Amer. Statist. Assoc.*, **95**, 1263-1268.

Sleasman, J. W., Nelson, R. P., Goodenow, M. M., Wilfert, D., Hutson, A., Bassler, M. Zuckerman, J., Pizzo, P. A. & Mueller, B. U. (1999). Immunoreconstitution After Ritonavir Therapy in Children With Human Immunodeficiency Virus Infection Involves Multiple Lymphocyte Lineages. *J. Pediatrics*, **134**, 597-606.